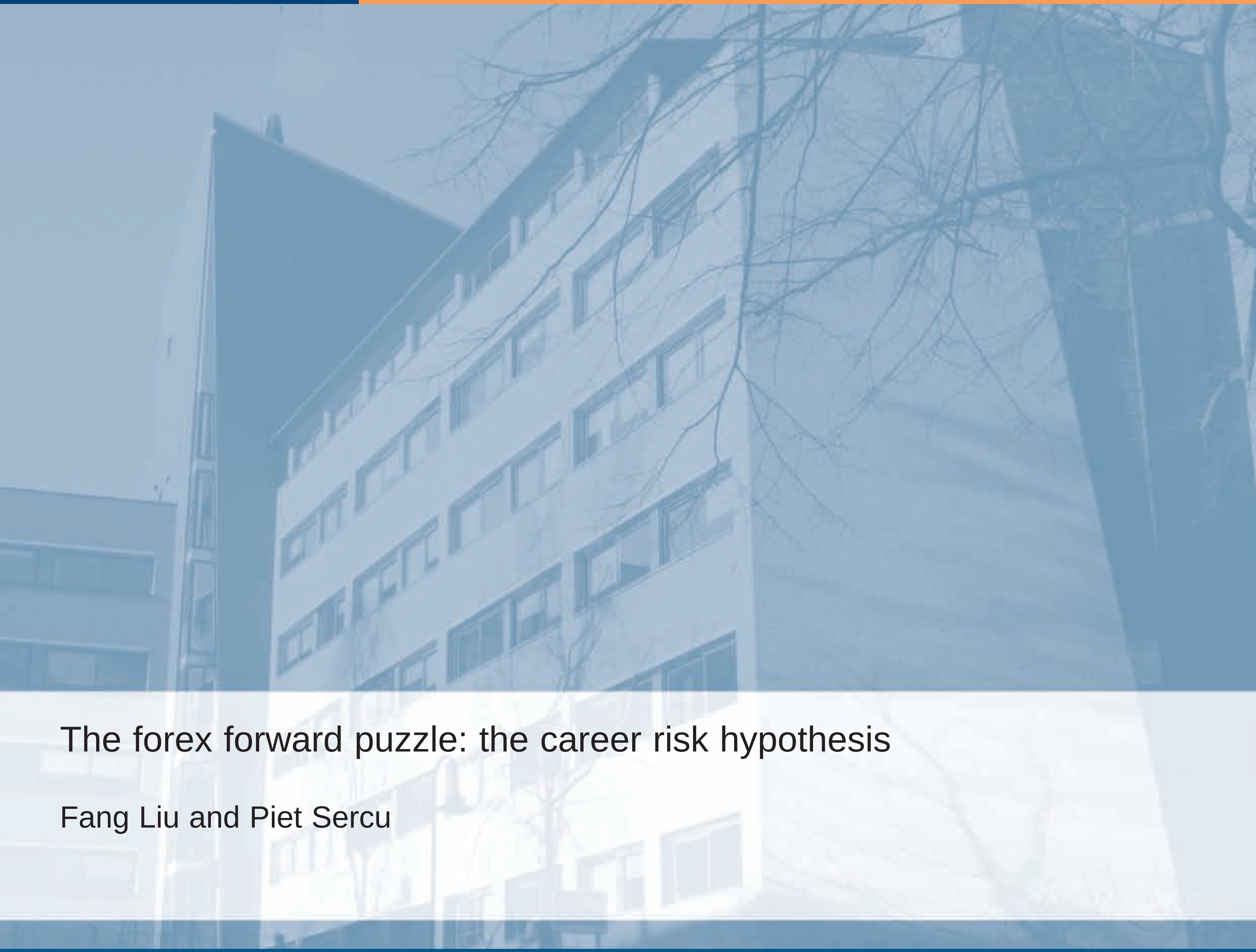
KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and Applied Economics



The forex forward puzzle: the career risk hypothesis

Fang Liu and Piet Sercu

DEPARTMENT OF ACCOUNTANCY, FINANCE AND INSURANCE (AFI)

AFI 0709

The Forex Forward Puzzle: the Career Risk Hypothesis

Fang Liu* and Piet Sercu†

First draft: July 2006
this version formatted at 14:47 on Monday 30th July, 2007

*KU Leuven, Naamsestraat 69, B- 3000 Leuven; email: fang.liu@econ.kuleuven.be.

†KU Leuven, Naamsestraat 69, B- 3000 Leuven; email: piet.sercu@econ.kuleuven.be.

Abstract

Prior work variously ascribes the forward puzzle—the low slope in the Fama (1984) regression of the exchange rate change on the forward premium—to various model misspecifications or statistical problems with non-stationary forward premia, but no single theory fully succeeds in explaining the puzzle. In this paper we simultaneously address the model-misspecification problem and the non-stationarity issue. On the basis of competing hypotheses about the risk premium we consider nonlinear models that all specify the Fama beta as approximately quadratic or spline functions of the forward premium. We estimate these relations using overlapping one-month observations for ERM-member exchange and forward rates against the DEM. The standard deviations are calculated under the Monte Carlo Method for overlapping observations. Wald test confirms the presence of such nonlinearities, and the models outperform the Fama in terms of various in the goodness-of-fit measures, but the spline adds little relative to the simple quadratic. To handle the non-stationary forward premium problem, we decompose the forward premium into a long-memory co-movement component and a short-term filtered forward premium. In regressions that link exchange-rate changes to the long-memory co-movement component the forward puzzle worsens, while it is substantially reduced when, instead, the filtered component is used as the regressor, suggesting that the filtered component loads relatively heavily on expectations and the slow-moving trend on the missing variable. Beta appears to be an inverse-U-shaped function of the forward premium. This contradicts the Bansal risk premium and the transaction-cost/limit-to-arbitrage hypotheses, but is consistent with a Fallen-angel effect, where traders or portfolio managers shun long positions in assets with danger signals like forward discounts.

Keywords: forward puzzle, uncovered interest parity, tests, Peso problem, risk premium
JEL-codes: G32, G34.

The Forex Forward Puzzle: the Career Risk Hypothesis

Introduction

One empirical puzzle in international finance is the size of the bias in the forward premium as a predictor of future exchange rate changes. The Unbiased Expectations Hypothesis (UEH) posits that, in the Fama regression of exchange rate changes on beginning-of-period forward premia, the slope should equal unity and the intercept zero. However, as shown by Cumby and Obstfeld (1984), Fama (1984), and many others after them,[1] the empirical coefficients are not only systematically below unity, but disconcertingly often even negative. The empirical results are all the more unexpected as, in unconditional tests over long periods, the cross section of time-series-average interest differentials does match the cross-section of time-series-average rates of appreciation quite well (Backus and Smith, 1993). One interpretation of the downward bias is that there is a missing variable that correlates with the forward premium, and a prime candidate is a risk premium.[2] Another view is that, because of the near-unit-root characteristics of the regressor, the usual confidence intervals are vastly understated, see Roll and Yan (2000). True, this would not explain why the Fama-regression betas tend to be so low; but it would at least stop us from taking negative betas so seriously.

Some of the above issues become even more puzzling if one considers exchange rates within the European Exchange Rate Mechanism (ERM). First, in these rates there is strong mean-reversion, that is, predictability; so the usual remark that interest rates do not predict because there is nothing to predict does not apply here (Sercu, Vandebroek and Wu, 2006). Second, in the ERM context also the long-memory property of forward premia is a real puzzle. Member countries coordinated their monetary policies, which should have led to co-movement in the interest rates, not randomly diverging rates. Also, it is hard to imagine that expectations of percentage changes in ERM exchange rates would be non-stationary. In light of this, the

[1] See Froot and Thaler (1990) for an early review; equally excellent and more recent surveys are by Hodrick (1987), Engel (1996), and Sarno and Taylor (2002).

[2] Frankel and Engel (1984), Domowitz and Hakkio (1985), Hodrick and Srivastava (1986); Hodrick (1987, 1989), Cumby (1987), Mark (1988), Engel (1996), Hollifield and Uppal (1997), Mark and Wu (1997), Backus, Foresi, and Telmer (2001), and Chinn and Frankel (2002) all fail to explain the forward puzzle well; but see Bansal (1997) for a dissident view.

most likely cause of near-nonstationarity in forward premia would then be the missing variable (like a risk premium), a conjecture we try to substantiate in this paper. A third reason for focusing on the ERM is that it had clear anchors: the official bilateral parities and the admissible band; the associated multilateral measure of health, the divergence indicator; and the interest differential against the DEM, which was even canonized as an EMU accession criterion later on. These elements are useful to test a new hypothesis about the forward bias, inspired by the fallen-angel effect in stock markets: bearing in mind their career prospects, portfolio managers shun assets that fell badly, in the recent past, or emit other very visible danger signals.

In the remainder of this intro we outline the paper's more technical ingredients: we study a non-linear relation, we filter out the long-memory component of the forward premium, and we take into account the near-unit-root problem in our significance tests.

Non-linearities have long antecedents as a potential explanation of the forward puzzle. Such a non-linearity can arise because the risk premium is approximately quadratic in the forward premium, as Bansal (1997) points out. The non-linearity could also be due to transaction costs[3] or "limits to arbitrage"[4], notably when traders ignore gains that are of insufficient size relative to risk or transaction costs. Lastly, the risk premium could come from a career-risk-premium effect, where any danger signal, like a pronounced forward discount, adds to the portfolio manager's reluctance to invest. Like the Bansal and transaction-cost models, the career-risk hypothesis proposes a particular nonlinear model that can be written as a Fama regression whose slope, beta, varies depending on the forward premium. In this paper we specify the Fama beta as a quadratic function or even a quadratic spline of the forward premium. We estimate these nonlinear models using overlapping one-month observations for EMU-member exchange and forward rates against the DEM. Wald test confirms the presence of the nonlinearities, and the models outperform the Fama version in terms of all standard goodness-of-fit measures, but the spline adds little relative to the simple cubic. The cubic model produces an inverse U-shaped beta plot, supporting the Fallen-angel hypothesis, as we shall see.

Our second ingredient, after non-linearity, is filtering. Filtering is inspired by the fact that, if forward premia are nonstationary, then this must come from either the expectations, or the risk premium (or, more generally, the missing variable), or both. In the ERM, expectations are

---

[3]Huisman *et al.* (1998), Obstfeld and Rogoff (2000).

[4]Lyons (2001), Villanueva (2005), Sarno, Valente and Leon (CEPR 2006), and Baillie and Kiliç (2006).

even less likely to be unit-root than in other data, so our money is on the risk premium. We decompose the time series of forward premia into a Hodrick-Presscot "trend" (which turns out to be non-stationary) and a stationary filtered component, and we re-run our generalized Fama regressions with either this long-run-memory component or its filtered part as the regressor. Consistently with the idea that the long-memory part is more closely related to the risk premium while the filtered component loads more heavily on the expectations, we see that betas for the filtered premia are much higher, while those of the "trend" component in the forward premium are clearly negative. In both betas, though, there is an inverted U shape, suggesting that the fallen-angel or career-risk factor has both long- and short-memory components.

As our third ingredient—handling non-stationarity—all our significance tests are based on Monte Carlo simulations where forward premia have strong memory and where, like in our data, the observations periods overlap. Thus, we unite Roll and Yan (2000) with Hansen and Hodrick's (1980) earlier correction for overlapping observations under OLS assumptions.

The structure of the paper is as follows. In Section 1 we review the non-linear models and our way of empirically implementing them. In Section 2 we describe the filtering, in Section 3 we come to the main results and in Section 4 conclude. The appendix explains the application of Monte Carlo Simulation.

1 Non-linear variants of the Fama regression

1.1 Some earlier non-linear models

Huisman *et al.* (1998) condition the Fama regression coefficients on the day-by-day cross-sectional variation of forward premia. They start from a particular view on where the missing variable comes from: friction in the market. Since real-world markets are subject to transaction costs, they argue, uncovered interest arbitrage cannot perfectly align expected exchange rates and forward premia. Most of the time, expectations of exchange rate changes are so small and diffuse that this friction-induced noise between expectations and premia largely obscures the theoretical parity between the two. However, there may be occasions where the market does expect unusually large changes; and if the impact of friction is essentially unaffected by the size of the expected change, then in these instances the signal-to-noise ratio must be relatively favorable. Highly positive or negative forward premia should therefore be better predictors than small premia. Cast in familiar statistical terms, the Fama regression suffers from an errors-in-the-regressor type bias towards zero, and for a given variance of the noise term this bias can be reduced by constructing a subsample where the variance of the regressor is larger. Huisman *et al.* test this model using panel techniques with a cross-currency constraint that ensures numeraire-invariance of the estimates. They report that large-variance observations generate Fama regression coefficients close to unity, and even substantially above unity if the definition of "large variance" is very strict.

Huisman *et al.*'s approach is similar, in spirit, to an earlier regression by Bilson (1981), who works with an equation of the type

$$\tilde{s}_{t,\Delta} = I_t \cdot [\alpha_1 + \beta_1 f_{t,\Delta}] + (1 - I_t) \cdot [\alpha_0 + \beta_0 f_{t,\Delta}] + \eta_{t,\Delta}, \tag{1.1}$$

where $\tilde{s}_{t,\Delta}$ is the percentage change, or log change, in the spot rate in the period $[t,\, t+\Delta]$; I_t is an indicator that the forward premium observed at t is among the n percent most extreme observations; and $f_{t,\Delta}$ is the percentage or log forward premium at t for delivery at $t+\Delta$. Thus, in both Huisman *et al.* and Bilson the Fama regression abruptly switches between parameters (α_1, β_1) and (α_0, β_0) depending on whether the forward observation is extreme. One can experiment with the criterion n — *e.g.* the five, ten or twenty percent biggest premia—and see which one works best.

The more recent Limits-to-Arbitrage literature comes up with transition functions that are, basically, smoother and more flexible variants of the Bilson equation. The Sarno *et al.* variant, for instance, would read like[5]

$$\tilde{s}_{t,\Delta} = [\alpha_1 + \beta_1 f_{t,\Delta}] + \Phi(f_{t,\Delta}, \gamma) \cdot [\alpha_2 + \beta_2 f_{t,\Delta}] + \eta_{t,\Delta}, \tag{1.2}$$

with $\Phi(f_{t,\Delta}) := 1 - \exp\left[-\gamma (f_{t,\Delta})^2\right]$. Φ is an inverse bell-shaped function that assigns zero weight to the (α_2, β_2) version of the regression when $|f|$ is zero, and almost unit weight when $|f|$ is very large. Accordingly, Sarno *et al.* hypothesize that while α_1 and β_1 may be close to zero, we still have $\alpha_1 + \alpha_2 = 0$ and $\beta_1 + \beta_2 = 1$.

Bansal (1997) takes a very different perspective, focusing on the risk premium instead of friction. He starts from an orthodox CCAPM equilibrium asset pricing model, and establishes that its currency risk premium is approximately quadratic in the forward premium. Thus, the entire relation between expected change and forward premium becomes quadratic—inverse U-shaped, to be more precise. In his tests, Bansal approximates this by a piecewise-linear,

[5] The argument in their Φ is the expectation, not the premium. Baillie and Kiliç (2006) do use the premium (in a logistic transition function).

Table 1: **Overview of the nonlinear models**

Model	rp in terms of f	β in terms of f
Huisman *et al.* (1998) and limit-to-arbitrage	—	U-shaped or inverse bell pattern
Bansal (1997)	– U-shaped pattern: – V-shaped approximation:	β negative linear in f $\beta > 0$ for $f < 0$; $\beta < 0$ for $f > 0$
Fallen-angel hypothesis	Cotangent shape, possibly asymmetric	inverse U or inverse V, possibly asymmetric

Key: "rp" denotes the risk premium, "β" the Fama (1983) beta and "f" the forward premium.

inverse V-shaped relation, with an output that is immediately interpretable:

$$\tilde{s}_{t,\Delta} = I_t^+ \cdot [\alpha_+ + \beta_+ f_{t,\Delta}] + (1 - I_t^+) \cdot [\alpha_- + \beta_- f_{t,\Delta}] + \eta_{t,\Delta}, \tag{1.3}$$

where I_t^+ is an indicator that the forward premium for period t is positive. Thus, in this model the Fama β changes discretely around $f = 0$; the hypothesis is that positive fs have a negative β and vice versa. If the approximation of the inverse U-shape by an inverse V is omitted, the Bansal equation can be written as involving a beta that, itself, is negative linear in f, thus producing the overall quadratic relation between the expectation and the forward premium. Table 1 sums up the models.

1.2 The Fallen-Angel Hypothesis

Our own tentative explanation is inspired by Sercu and Vinaimont's (2006) work on Peso risk, which seemed a promising but ultimately unsuccessful candidate explanation for the forward bias in the private ECU. Like the original Peso hypothesis, our hypothesis invokes "dark matter", risks not observed by the econometrician. In the original Peso version, the dark matter is a low-probability, huge change. The potential change being huge, it does affect the expectation and therefore the forward premium; but in a finite sample the low-probability change may never be observed, so that the statistician would conclude that the forward premium systematically mispredicts the future spot rate.[6] One problem with this view

[6] For this to affect the regression coefficient rather than the intercept, the Peso risk must be time-varying and correlated with the forward premium. A plausible mechanism is as follows. When bad news about the foreign currency hits the market, the spot rate drops. But the concomitant selling of short-term paper (or borrowing against deposits) also pushes up the foreign interest rate, thus seemingly foretelling a further drop—or, if you

is that most of the empirical evidence comes from floating rates, and one wonders what the huge Peso event might be if there is no system of interventions or exchange restrictions that keep the accumulating tensions bottled up for longish times. If one accordingly rejects the Peso view as implausible for floating rates, then it may seem that we are only inches away from the overreaction hypothesis. In this view, the huge change fails to materialize not because its probability is low, but because it exists only in the minds of the traders. People are subject to bouts of panic or overoptimism, causing soon-corrected movements in spot rates accompanied by changes in interest rates in the opposite direction. This fads & fashions view is what Sercu and Vinaimont ultimately come to for the private ECU.

Our own dark-matter variant, in contrast, invokes no such irrationality. The starting point is that the market is dominated by professional investors (traders or portfolio managers), not individuals playing with their own stakes. For a professional, the ultimate decision criterion is the portfolio manager's career and remuneration prospects. This is not the same as the return on the portfolio to be managed because PV-ed remunerations and reputation are not linear in the portfolio return, and depend also on how and when any losses have occurred. Imagine, again, bad news about a foreign currency, immediately showing up in a falling spot rate and a falling forward premium (rising foreign money-market rates). The manager may play it safe and liquidate the foreign positions, thus risking to miss a recovery; or she may act contrarian and stay long, risking a further drop in the spot rate. In making the choice she will note that a cash loss looks worse than an opportunity loss, in general. But a cash loss from being contrarian (when there has been a clear and publicly observable bad initial signal) looks much worse than an opportunity loss from missing a rally which, judging by the initial forward premium, was deemed to be rather unlikely anyway. Any cash loss from going against the flow will be met with the comment that the trader "should have seen it coming", but the opportunity loss from following the consensus signal will not. In short, when bad news hits the market, professional investors head for the exit even if there is an expected gain from the subsequent recovery, because the expected gain from the recovery is counterbalanced by a dark matter, the potential damage to the professional investor's career if expectations turn out to be wrong. In the stock market this is known as the "fallen angel" effect: stocks that did badly are shunned by portfolio managers and, therefore, generate high returns.[7]

---

wish, slowing down the immediate drop. If the Peso event then fails to materialize, the peaking forward premium tends to be followed by a recovery in the spot rate, producing the negative regression coefficients.

[7]See Ikenberry, D., Lakonishok, J. and Vermaelen.T. (1995).

This particular dark-matter theory is testable, unlike the strict Peso view or its (statistically indistinguishable) overreaction counterpart, because it does not involve invisibles. Instead, it predicts a risk premium for holding currencies with public danger signals. An unusually negative forward premium would certainly be one such warning light, triggering a positive extra required return. The amount of expected return the manager is willing to give up is the career risk premium, which is expected to be positive and large when the forward premium is negative and to be unimportant when the forward premium is around zero. Similarly, the manager would be willing to give up some expected return for the safety promised by a markedly positive forward premium. In short, the missing variable exhibits a cotangent-shaped relation to f. Note that the relation between the forward premium and the fallen-angel premium needs not be symmetric. While it takes a large bribe to go against a warning signal and risk a cash loss, a smaller return shortfall may already be enough to make the trader ignore a positive signal, because the risk of ignoring the signal is just an opportunity cost.[8]

Thus, to model the private risk premium as a function of f we chose a negative-sloping and possibly asymmetric function that is probably quite flat when f is close to zero but raises faster for larger or for more negative values of the forward premium. Let us use the notation $x_+ := \mathrm{Max}(x, 0)$ and $x_- := \mathrm{Min}(x, 0)$ to denote observations selected by sign. Within the class of low-order polynomials we could then chose a piecewise quadratic, like $\zeta\ f_-^2 - \eta\ f_+^2$ with $\zeta > 0$, $\eta > 0$, or a piecewise cubic like $-\zeta\ f_-^3 - \eta\ f_+^3$. The corresponding test equations are,

$$
\begin{aligned}
&\text{fallen-angel premium} && \\
\text{(quadratic:) } \mathrm{E}_t(\tilde{s}_{t,\Delta}) - f_{t,\Delta} &\approx \zeta\,(f_{t,\Delta})_-^2 - \eta\,(f_{t,\Delta})_+^2; \\
\Rightarrow \mathrm{E}_t(\tilde{s}_{t,\Delta}) &\approx [1 + \zeta\,(f_{t,\Delta})_- - \eta\,(f_{t,\Delta})_+]\, f_{t,\Delta},
\end{aligned} \tag{1.4}
$$

$$
\begin{aligned}
\text{(cubic:) } \mathrm{E}_t(\tilde{s}_{t,\Delta}) - f_{t,\Delta} &\approx -\zeta\,(f_{t,\Delta})_-^3 - \eta\,(f_{t,\Delta})_+^3; \\
\Rightarrow \mathrm{E}_t(\tilde{s}_{t,\Delta}) &\approx [1 - \zeta\,(f_{t,\Delta})_-^2 - \eta\,(f_{t,\Delta})_+^2]\, f_{t,\Delta}.
\end{aligned} \tag{1.5}
$$

Thus, here the betas are predicted to be inverted U- or V-functions of f.

While it is not difficult to test whether the predicted patterns are present or not, finding that they are does not necessarily mean that they reflect a fallen-angel effect. Thus, before starting the main tests we want to verify whether it is generally true that danger signals

---

[8] This would be even more so if the foreign currency is exotic and the home currency a major one. True, the international market is a mix of various nationalities; and if the pattern is as described above, then we have the opposite pattern if we change numeraire. But it is also a fact that more managers report in, say, GBP or USD than in DKK or BEF; thus, if there is an asymmetry in the private risk premia, the bigger currency's point of view is likely to dominate.

other than the forward premium seem to generate a fallen-angel risk premium. If so, this finding would lend extra credibility to our interpretation that also a big forward premium is a fallen-angel signal.

1.3 Preliminary Tests: do Danger Signals Lead to Excess Returns?

The currencies we work with are the Belgian Franc (BEF), German Mark (DEM), Danish Krone (DKK), French Franc (FRF), Dutch Guilder (NLG), Spanish Peseta (ESP), Irish Punt (IEP), Italian Lira (ITL) and Austrian Schilling (ATS). All data are acquired from DataStream. We use weekly observations on one-month forward contracts, and the sample period is from January 1st, 1976 to December 31st, 1998 (1200 observations). The one exception is the IEP, whose Datastream coverage starts on April 2nd, 1979 (1030 observations). The future spot rate is the spot rate on the delivery day, two working days plus 30 days after the date when the transaction is agreed. DEM works as the base currency; that is, exchange rates equal the value of one DEM in units of the other currency.

Let us define the excess return, $R^e_{t,\Delta}$, as the exchange-rate change in excess of the forward premium. Under the UEH the expected excess return should be zero in absence of the risk premium and irrational expectations. In the modified Fama regression,

$$\begin{aligned} R^e_{t,\Delta} &:= \tilde{s}_{t,\Delta} - f_{t,\Delta} \\ &= \alpha + (\beta - 1) f_{t,\Delta} + \sum_j \gamma_j X_{j,t} + \eta_{t,\Delta}, \end{aligned} \tag{1.6}$$

the intercept α and slope $(\beta-1)$ should both equal zero under the null hypothesis. In addition, extra regressors X_j should have no explanatory power. In this section we verify whether excess returns load positively on danger signals other than the forward premium, as our fallen-angel hypothesis predicts. As possible proxies for danger signals, we select the following seven variables related to the divergence indicator and to recent trends in the spot rate over one day:

1. Position in the ERM band

   The European Exchange Rate Mechanism (ERM) was built around the ECU, a basket of all EU currencies. For each currency there was a target value in the basket, called central parity. Whenever the actual value of the ECU moved too far from the central parity, the member state had to take "policy measures" to bring back its currency into line. The divergence indicator (d) provides the signal. It is calculated as the divergence between the actual value and central parity of the ECU, in units of home currency as a percentage of the allowed maximum divergence,[9]

$$d := \frac{\left[\text{actual value} - \text{central parity}\right]/\text{central parity}}{\text{maximum divergence}} \tag{1.7}$$

A positive d means a strong ECU, that is, a weak home currency and therefore, under the career-risk hypothesis, a risk premium for holding it. Since the exchange-rate data in the tests are units of home currency per DEM, the career-risk risk premium on home currency translates into a negative risk premium for holding DEM. That is, the relation between our risk premium and the d should be negative.

The longer the time the divergence indicators has been positive, the worse the signal. Two versions of the d are considered that measure the persistence of the d: weekly and monthly averaged divergence indicators, denoted as $\overline{d}_{w,t}$ and $\overline{d}_{m,t}$.

$$\begin{aligned} \overline{d}_{w,t} &:= \sum_{l=1}^{7} d_{t-l}/7 \\ \overline{d}_{m,t} &:= \sum_{l=1}^{30} d_{t-l}/30 \end{aligned} \tag{1.8}$$

2. Change of the position within the band

   A weakening of home currency against the ECU could be another danger signal, over and above the level of the d. We look at the change over the last 24 hours:

$$d_{t,-1} = d_t - d_{t-1} \tag{1.9}$$

   A positive change means a strengthening ECU, that is, a weakening home currency. Like a positive level of the ECU, a positive change should therefore get a negative sign in the risk premium on the DEM.

3. The change of the bilateral exchange rate for the DEM.

   The variables in the first two groups are in terms of the basket, the ECU. The d is, however, a relatively weak constraint in the sense that it looks at the average deviation of the member currencies from their central parities, not the highest pairwise deviation; and it may trigger “policy measures” like interest-rate changes but not intervention in the exchange market. Such intervention is based on the bilateral rates, which should stay

[9]The maximum divergence depends on the currency’s weight in the ECU. Against the DEM, for example, the ECU cannot drop as far as against the IEP, since 30 percent of the the ECU is DEM and only 1 percent is IEP.

Table 2: **Weights of the Risk Variables in the Principal Components ξ_1, ξ_2 and ξ_3, and Explanatory Power**

$$\xi_1 = \kappa_{11}d + \kappa_{12}\overline{d}_{w,t} + \kappa_{13}\overline{d}_{m,t} + \kappa_{14}s_{t,-30} + \kappa_{15}s_{t,-1} + \kappa_{16}s_{t,-7} + \kappa_{17}d_{t,-1}$$
$$\xi_2 = \kappa_{21}d + \kappa_{22}\overline{d}_{w,t} + \kappa_{23}\overline{d}_{m,t} + \kappa_{24}s_{t,-30} + \kappa_{25}s_{t,-1} + \kappa_{26}s_{t,-7} + \kappa_{27}d_{t,-1}$$
$$\xi_3 = \kappa_{31}d + \kappa_{32}\overline{d}_{w,t} + \kappa_{33}\overline{d}_{m,t} + \kappa_{34}s_{t,-30} + \kappa_{35}s_{t,-1} + \kappa_{36}s_{t,-7} + \kappa_{37}d_{t,-1}$$

		weights κ							
		d	$\overline{d}_{w,t}$	$\overline{d}_{m,t}$	$s_{t,-1}$	$s_{t,-7}$	$s_{t,-30}$	$d_{t,-1}$	variance explained
	ξ_1	-0.58	-0.58	-0.58	-0.02	-0.02	-0.03	-0.03	0.43
BEF	ξ_2	0.02	0.03	0.03	-0.53	-0.62	-0.56	-0.14	0.26
	ξ_3	-0.05	0.04	0.03	-0.09	0.15	0.18	-0.97	0.14
	ξ_1	-0.58	-0.58	-0.58	0.02	0.03	0.01	-0.02	0.42
DKK	ξ_2	-0.04	-0.00	-0.01	-0.53	-0.61	-0.53	-0.24	0.26
	ξ_3	-0.04	0.05	0.02	-0.11	0.17	0.33	-0.92	0.14
	ξ_1	-0.57	-0.57	-0.57	0.05	0.09	0.14	-0.02	0.43
FRF	ξ_2	-0.09	-0.10	-0.09	-0.52	-0.64	-0.54	-0.03	0.25
	ξ_3	-0.04	0.06	0.04	-0.11	0.06	0.08	-0.99	0.14
	ξ_1	-0.58	-0.58	-0.57	-0.01	-0.02	-0.06	-0.03	0.42
NLG	ξ_2	0.02	0.04	0.04	-0.54	-0.61	-0.55	-0.17	0.26
	ξ_3	-0.06	0.06	0.03	0.07	0.11	0.12	-0.98	0.14
	ξ_1	-0.57	-0.57	-0.57	-0.05	-0.08	-0.10	-0.02	0.43
ITL	ξ_2	0.08	0.08	0.08	-0.53	-0.63	-0.55	-0.04	0.26
	ξ_3	-0.04	0.05	0.03	0.15	-0.05	0.00	-0.99	0.14
	ξ_1	-0.58	-0.58	-0.57	0.02	0.04	0.05	-0.03	0.42
IEP	ξ_2	-0.03	-0.04	-0.04	-0.52	-0.64	-0.57	-0.01	0.26
	ξ_3	-0.06	0.07	0.04	0.03	0.00	-0.01	-0.99	0.14
	ξ_1	-0.50	-0.49	-0.49	0.17	0.28	0.39	-0.09	0.51
ESP	ξ_2	-0.14	-0.26	-0.24	-0.63	-0.38	-0.14	0.54	0.23
	ξ_3	-0.21	-0.15	-0.18	0.10	-0.58	-0.43	-0.61	0.13

within the ±2.25% band. We look at the bilateral rate that is most likely to be troublesome, for weak currencies: the value in the DEM. Changes in that rate are measured in three time spans: daily, ($s_{t,-1}$), weekly ($s_{t,-7}$), and monthly ($s_{t,-30}$).

$$\begin{aligned} s_{t,-1} &= \ln S_t - \ln S_{t-1} \\ s_{t,-7} &= \ln S_t - \ln S_{t-7} \\ s_{t,-30} &= \ln S_t - \ln S_{t-30} \end{aligned} \tag{1.10}$$

A rise in the DEM's value means a danger signal and, therefore, a lower risk premium for the DEM.

In short, under the career-risk view all seven potential danger signals should be negatively correlated with the risk premium on the DEM.

Principal Component Analysis (PCA) was used to compress the seven dimensions into three principal components. The empirical result of our PCA is exhibited in Table 2. The first

Table 3: **The Predictive Power of the Principal Components**

$R^e_t = \alpha + \gamma_1\xi_{1,t} + \gamma_2\xi_{2,t} + \gamma_3\xi_{3,t} + \upsilon_t$

	γ_1	γ_2	γ_3	R^2
BEF	***0.0008	***0.0003	***0.0003	0.024
DKK	***0.0005	0.0001	*0.0002	0.006
FRF	***0.0003	***0.0005	-0.0001	0.008
NLG	***0.0006	***0.0004	***-0.0002	0.078
ITL	***-0.0008	**-0.0003	-0.0002	0.009
IEP	***-0.0003	*0.0002	0.0003	0.002
ESP	***0.0007	0.0003	0.0004	0.002
number of right signs	5	6	4	

principal component ξ_1 can be summarized as the position factor, for in this linear combination the variables d, $\overline{d}_{w,t}$ and $\overline{d}_{m,t}$ have dominant weights relative to the other variables. The second component ξ_2 can be viewed as the bilateral change factor because of the prominent weights for $s_{t,-1}$, $s_{t,-7}$ and $s_{t,-30}$. The third component ξ_3 relates to the change in the position, primarily weighted by $d_{t,-1}$. Since the main variables have negative weights in their eigenvector, a higher value of the principal component should be associated with a higher expected return on the DEM; that is, the regression coefficient should be positive.

Now, we test if our conjecture is right or not by regressing the excess return on the principal components in Equation (1.11),

$$R^e_t = \alpha + \gamma_1\xi_{1,t} + \gamma_2\xi_{2,t} + \gamma_3\xi_{3,t} + \upsilon_t. \tag{1.11}$$

The estimations and the goodness-of-fit of the models are reported in Table 3. While under the null no coefficient should be significant, bar perhaps one or two on a pure-chance basis, we see fifteen starred estimates. The intervention factor ξ_1 has five times a significant positive coefficient (the exceptions being the currencies with large forward premia, ITL and IEP). ξ_2 comes up with a positive value in six equations, albeit with unsatisfactory significance in two cases. The third component, ξ_3, has a positive coefficient in four equations, although only two of them are significant. Most of the wrong signs are insignificantly different from zero. For each currency, the R^2 is low but is nevertheless much higher than for the regression where the forward premium is the regressor. So, the fallen-angel hypothesis acquires some credibility: the empirical evidence implies that general danger signals other than the interest differential trigger higher expected returns, consistent with the idea that career- or image-risk considerations drive the manager to ask for extra return. We now turn to the forward premium to see whether it seems to as a danger signal of its own.

1.4 Main Test Equation

Our test models are part of the literature that focuses on the risk premium, which explains the puzzle as a failure to allow for a risk premium in the Fama regression. Generalizing Bansal's approach, we present two models that approximately fit all of the above hypotheses. The first one is a cubic model where the risk premium takes the form of higher orders of the forward premia,

$$\begin{aligned} \mathrm{E}_t(\tilde{s}_{t,\Delta}) &= \alpha + \beta_1 f_{t,\Delta} + \beta_2 f_{t,\Delta}^2 + \beta_3 f_{t,\Delta}^3 \\ &= \alpha + f_{t,\Delta} + \underbrace{(\beta_1 - 1) f_{t,\Delta} + \beta_2 f_{t,\Delta}^2 + \beta_3 f_{t,\Delta}^3}_{\text{risk premium}}, \end{aligned} \tag{1.12}$$

Thus, the Fama beta becomes quadratic in the forward premia in equation (1.13),

$$\beta(f_{t,\Delta}) = \beta_1 + \beta_2 f_{t,\Delta} + \beta_3 f_{t,\Delta}^2. \tag{1.13}$$

Section 3 examines the nonlinearity by testing the joint hypothesis $\beta_2 = \beta_3 = 0$. If and when the null of a linear model is rejected, we can check the observed pattern against the alternatives set forth in Table 1. But the quadratic approximate Fama β may lack flexibility; the tails of an inverse bell shape, for instance, can not be captured. As a more flexible alternative to the cubic model in Equation (1.12), we therefore let the Fama beta be a quadratic spline function of the forward premia:

$$\beta(f_{t,\Delta}) = \beta_1 + \beta_2 f_{t,\Delta} + \beta_3 f_{t,\Delta}^2 + d_1 (f_{t,\Delta} - k_1)_+^2 + \ldots + d_p (f_{t,\Delta} - k_p)_+^2; \tag{1.14}$$

In this equation the Fama regression is implicitly revised into a nonlinear form,

$$\mathrm{E}_t(\tilde{s}_{t,\Delta}) = \alpha + [\beta_1 + \beta_2 f_{t,\Delta} + \beta_3 f_{t,\Delta}^2 + d_1 (f_{t,\Delta} - k_1)_+^2 + \ldots + d_p (f_{t,\Delta} - k_p)_+^2] f_{t,\Delta} + \eta_t. \tag{1.15}$$

Familiarly, the quadratic spline function in equation (1.14) consists of two parts: a quadratic function of f and the plus functions, $(f - k_i)_+^2 := [\mathrm{Max}(f - k_i, 0)]^2$ for $i = 1, ..., p$. The pre-set parameters k_i are referred to as the knot points; so the squared positive differences between f and the knot points are included into the plus functions. Like a $(p+3)$-th degree polynomial, the spline function is continuous in its level and first-order derivative, but it allows the second-order derivative to change at each knot point without any repercussions on the function for lower values of f. Here, we set the knot points as follows. The percentage forward premia are ranked by size, and for each currency six knot points separate the whole set of observations into seven bands. The values of the knots for a given currency are the 5th, 10th, 20th, 80th, 90th and 95th top percentile values of the sample of the forward premia of the currency. Note

Table 4: **Knot-point values of frf for the spline**

variable (%)	Minimum	knot1	knot2	knot3	knot4	knot5	knot6	Maximum	Mean	S.D.
Cum prob (%)	0	5	10	20	80	90	95	100		
f	-0.15	0.00	0.01	0.04	0.53	0.69	0.88	2.61	0.33	0.33
$\widehat{f}$	-0.05	0.06	0.08	0.10	0.53	0.63	0.69	0.98	0.33	0.21
ff	-0.52	-0.28	-0.22	-0.12	0.09	0.21	0.39	2.08	0.00	0.25

Key: "f" is the forward premium, "$\widehat{f}$" is the co-movement component," ff" is the filtered forward premium, "Cum prob" is the cumulative probability and "S.D." is the standard deviation. All the numbers are expressed in percent.

that the first band corresponds to the lowest premia, and the seventh band to the highest premia. While the central zone is wide in terms of frequencies, in algebraical terms it is not wide because the density is much higher in the middle. For example, in Table 4 we show the knot points for the raw forward premium of the FRF, as well as for its two decompositions to be introduced in Section 2.

Empirical results are reported in Section 3, but we first explain how the filtering and the significance testing is done.

2 Filtering Out the Long-memory Co-movement in the Forward Premia

In this section we consider the near-nonstationarity of the forward premium. Non-stationarity in a variable invalidates the standard statistical model and makes the usual t-statistics unreliable. Regression, in such a framework, makes sense only if the dependent variable co-integrates with the independent variable with at least one co-integration vector. In Table 5 we show the results of the Fractional Integration test (FI)[10] to examine the non-stationary nature of the variables on both sides of the Fama regression.

In Table 5, unsurprisingly, the exchange rate changes $\tilde{s}_{t,\Delta}$ are classified as stationary with $d < 0.5$ in all currencies. The forward premia $f_{t,\Delta}$, in contrast, seem non-stationary except those for the ITL. So, the traditional regression has an imbalance problem.

Roll and Yan (2000) speculate on the likely source of the non-stationarity of the forward

[10]In the analysis of persistency of the time series, we usually consider the order of integration, d. When d is smaller than 0.5, the process is (weakly) stationary with finite variance. If d equals 0.5 or greater, the process is non-stationary with infinite variance. Infinite variance is more general than the unit-root case ($d = 1$). Here, we estimate d with Wavelet Ordinary Least Square (WOLS), See Tkacz(2001).

Table 5: **Fractional Integration Test**

d	$\tilde{s}_{t,\Delta}$	$f_{t,\Delta}$	ff	ω^*
ATS	0.109	0.544	-0.098	—
BEF	0.299	0.507	-0.312	—
DKK	0.257	0.608	-0.317	—
FRF	0.235	0.572	-0.201	—
NLG	0.109	0.561	-0.043	—
ESP	0.239	0.564	-0.294	—
IEP	0.184	0.581	-0.363	—
ITL	0.205	0.444	-0.310	—
ω^*	—	—	—	***-0.84

s denotes the spot return, f_Δ the forward premium for horizon Δ, ff its H-P-filtered version, and ω^* is the estimated common trend in f. If d is larger than 0.5, the diagnosis is that the time series looks non-stationary, otherwise, it is viewed as stationary.
.

premium. They conceptually decompose the nominal interest rate differential into the real interest rate differential and the expected inflation differential. The real interest rate differential might be a unit-root process when there is an economically vast gap between countries, but that is not the case for the mainstream economies studied in most published tests. So, they conjecture that the persistence in the forward premium is due to a non-stationary inflation risk premium or a non-stationary expected inflation differential.

An alternative approach is to conceptually dissect the forward premium into the expected spot-rate change and the missing variable, possibly a currency-risk premium. In principle, any non-stationary expectations for exchange-rate changes should show up in the Fractional Integration tests on the realized changes. True, in finite samples there is an issue of power here: the variability of even a non-stationary expectation may still be small relative to the white-noise component in $\tilde{s}_{t,\Delta}$ and, therefore, hard to establish beyond the usual doubt. In our case, there is a good *a priori* argument, though: it is hard to imagine non-stationary expectations for percentage changes in exchange rates that all belong to the ERM.

For the missing variable, the plausibility of non-stationarity is harder to gauge *a priori* as we do not even know what the missing variable is. But, in a way, we do not need priors: if observed forward premia are non-stationary or close to it and if we reject unit roots for the expectations part, then the non-stationarity must come from the residual, the missing variable. Being likely to be a long-run process in time domain or a low-frequency noise in frequency domain, we estimate the long-memory component on the basis of the Hodrick-Prescott trends in each forward premia series. The Hodrick-Prescott (HP) filter is a standard instrument to capture the trend in time series. The HP-filter works as follows. A series f_j for currency j is to be decomposed into "trend" and "cycle" components,

$$f_{j,t} = f_{j,t}^{tr} + f_{j,t}^{c}. \tag{2.16}$$

The HP filter estimates the trend from the solution to the following minimization problem with a pre-set smoothing parameter λ:

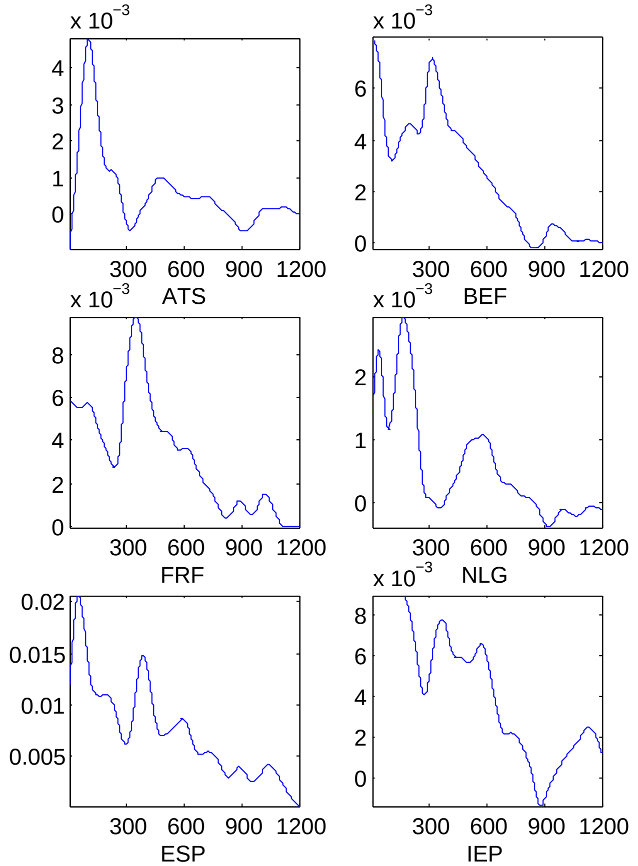
$$\min_{f_{j,t}^{tr},\forall t} \sum_{t=1}^{T} (f_{j,t} - f_{j,t}^{tr})^2 + \lambda \sum_{t=1}^{T} \left[\underbrace{(f_{j,t+1}^{tr} - f_{j,t}^{tr}) - (f_{j,t}^{tr} - f_{j,t-1}^{tr})}_{\text{second difference in the trend}} \right]^2 . \tag{2.17}$$

This objective function consists of two parts: the squared sum of the deviations of the series from its trend, summing up the badness-of-fit of the trend series vis-a-vis the raw data; and the sum of the squared trend reversals (the difference between two consecutive first differences), measuring the smoothness of the trend series. λ, the smoothing parameter, weighs the components in the objective. For quarterly data, for instance, it is conventionally set at 1600; we use the default value in Eviews 5.0 for weekly data, $\lambda = 270400$. Applying this procedure for each forward premia series, we obtain a currency-specific trend series f_j^{tr} for every currency j. Figure 1 shows the eight trends. A strong similarity of the patterns is immediately obvious: forward premia against DEM are generally falling over time, in line with the general level of interest rates and, at the end, converging with the approach of a single monetary policy under the authority of the ECB. The existence of important links between the eight trend series is confirmed when we compute correlation coefficients. Panel B of Table 6 shows the correlations between the eight HP trend components. Most of the values are above 0.60, much higher than the correlation coefficients between the raw forward premia in Panel A. Thus, series-by-series trends that consist of the sluggish (or, tentatively, long-memory) components of the premia are behaving quite similarly, suggesting that they load more heavily on one underlying common trend.

Thus, instead of working with the individual trends we can assume that here is a common trend, ω_t, underlying the individual trend variables. The motivation for using one common trend is that the individual trends may be over-fitting their series: by retaining only the common part, an additional filter is administered. Our proxy for the common component ω_t is calculated as follows. We first scale each trend series f_j^{tr} by its mean, $\bar{f}_j$, and we then average across the currencies j:

$$\omega_t := \frac{\sum_{j=1}^{J} \frac{f_{j,t}^{tr}}{\bar{f}_j}}{J}. \tag{2.18}$$

Figure 1: **the HP trends in individual currencies**



An issue worth raising here is that, because our data set has a small cross-section (only eight currencies), we still have to consider potential distortion from extreme observations. For instance, we may have missed some of the data errors in Datastream. A standard solution is to truncate the input data. So in our calculations, following the normal distribution law, only the values within the two-standard-deviation interval around the medians are included in the average calculations.

$$\omega_t^* := \frac{\sum_{j=1}^{J} I_{j,t}^{2\sigma} \frac{f_{j,t}^{tr}}{f_j}}{\sum_{j=1}^{J} I_{j,t}^{2\sigma}},. \tag{2.19}$$

where $I_{j,t}^{2\sigma}$ indicates whether the (j,t)-th trend observation is within two standard deviations from its median. So, when the value is outside the two-σ interval, it has no weight in the average.

The better the individual forward premia f_j correlate with the common trend ω^*, the more co-movement it captures. Panel D of Table 6 shows high correlation coefficients between these two variables, from 0.61 to 0.96, indicating that there are lots of commonalities and that ω_t^* works well as a proxy for the true co-movement. ω_t^* also has long memory: the unit root can

Table 6: **Correlations of f or components (eight currencies against dem)**

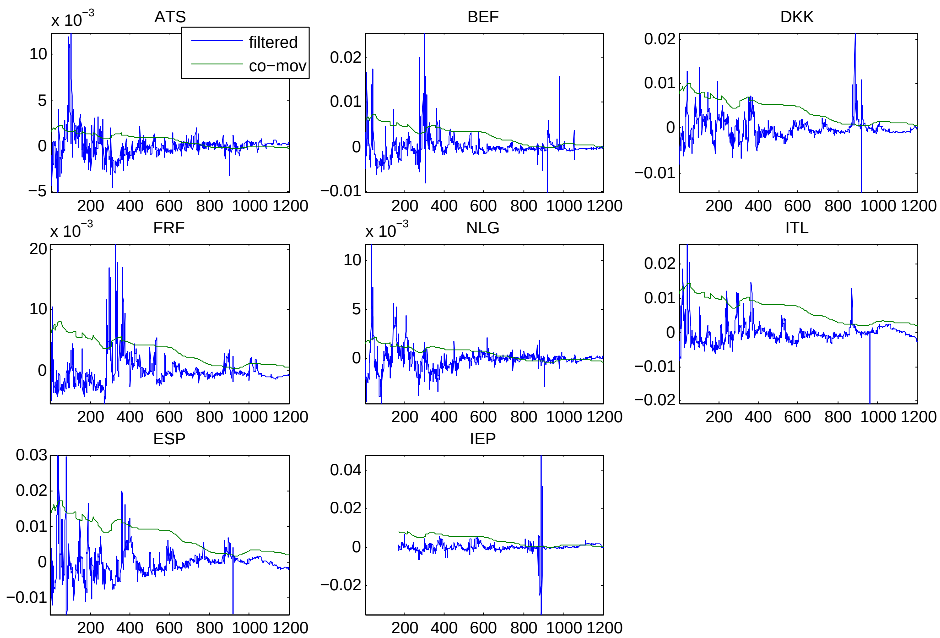
	ATS	BEF	DKK	FRF	NLG	ESP	IEP	ITL
Panel A: Raw forward premia, corr(f_j, f_k)								
ATS	1							
BEF	0.139	1						
DKK	0.328	0.561	1					
FRF	0.111	0.681	0.611	1				
NLG	0.551	0.313	0.420	0.179	1			
ESP	0.304	0.455	0.657	0.557	0.362	1		
IEP	0.277	0.469	0.393	0.453	0.387	0.449	1	
ITL	0.269	0.684	0.602	0.792	0.266	0.601	0.450	1
Panel B: HP trends, corr(f_j^{tr}, f_k^{tr})								
ATS	1							
BEF	0.264	1						
DKK	0.736	0.742	1					
FRF	0.293	0.897	0.739	1				
NLG	0.666	0.548	0.767	0.354	1			
ESP	0.582	0.790	0.887	0.767	0.682	1		
IEP	0.626	0.860	0.826	0.789	0.672	0.858	1	
ITL	0.420	0.913	0.795	0.860	0.547	0.908	0.827	1
Panel C: HP-filtered premia, corr($f\!f_j, f\!f_k$)								
ATS	1							
BEF	0.118	1						
DKK	0.257	0.395	1					
FRF	0.131	0.302	0.346	1				
NLG	0.150	0.357	0.298	0.039	1			
ESP	0.065	0.220	0.269	0.044	0.380	1		
IEP	0.065	0.071	-0.065	0.058	0.067	-0.027	1	
ITL	0.227	0.340	0.351	0.367	0.189	0.128	0.003	1
Panel D: corr(ω^*, f_j)								
correlation	ATS	BEF	DKK	FRF	NLG	ESP	IEP	ITL
coefficient	0.61	0.89	0.91	0.82	0.78	0.93	0.96	0.92

not be rejected by Dicky Fuller test shown in Table 5. Having obtained a single common-trend variable for all currencies under the HP filter method, we lastly decompose the eight series of f_j into a part that is linearly related to the common trend, and an orthogonal component:

$$\begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_J \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_J \end{bmatrix} + \begin{bmatrix} \phi_1 \\ \phi_2 \\ \dots \\ \phi_J \end{bmatrix} \times \omega_t^* + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_J \end{bmatrix}. \tag{2.20}$$

Below, the fitted values $\widehat{f}_{j,t} := \mu_j + \phi_j \omega_t$ are referred to as the co-movement component; and the retrieved residuals ϵ_j are henceforth referred to as the filtered forward premia, denoted as $f\!f_j$, and expressed as $f\!f_{j,t} = f_{j,t} - \widehat{f}_{j,t}$. Figure 2 shows graphs of the filtered premia, for

Figure 2: **The filtered forward premia and their trends**



comparison, along with the time series plots of the co-movement components $\widehat{f}_j$ underlying ω in Figure 1. By design, while for the co-movement components—and, by implication, the common trend—there is a lot of smoothness and inertia, in Figure 2 the random component is quite strong. A second difference is that most of time the plots of the filtered premia are below the plots of the co-movement components. Third, the co-movement component $\widehat{f}_j$ has a narrower standard deviation than the filtered forward premium ff_j. For the FRF, this can also be seen from the standard deviations of f, $\widehat{f}$ and ff in Table 4.

If our inference that the co-movement component $\widehat{f}_j$ catches the nonstationarity of f_j is right, its complement, the filtered premium $\mathit{ff}_{j,t}$, may be stationary. Table 5 compares the Fractional Integration parameter d of the filtered forward premia with the raw forward premia. In each currency $\mathit{ff}_{j,t}$ has a d smaller than 0.5, indicating that we seem to have a stationary series after filtering out the co-movement component from f. At the same time, the filtered premia should also be more independent from each other if the common part does the job it is expected to do. From Panel C of Table 6, the correlations between the ffs are much lower than the correlations between the fs, which confirms that the co-movement component captures most of the correlations in the forward premia.

The fact that the co-movement component is strongly long-memory leads to the conjecture that it contains little or no forecasts of exchange-rate changes; that is, the presence of this common trend may have decreased the predictive power of the forward premium—if at least the other component, the filtered part, does load more heavily on the expectations and not just on the transient part of the risk premium. This is what the tests of Section 3 are about.

3 Empirical Results

We start with the non-linear model as applied to the non-filtered forward premia (Sections 3.1 and 3.2), and then we work their common-trend or the filtered components (Section 3.3).

3.1 Nonlinearity Test and Model Evaluation

Table 7 shows the summary results for the linear, cubic and spline regressions. The values for the standard Fama β in the second column are significant and positive, a rare result in this literature. However, we still find the Fama model is misspecified: there is a clear non-linear relation between the spot changes and the forward premia. Here, nonlinearity is demonstrated if the Wald test rejects the null hypothesis of a linear relation, the original Fama model; in other words, we jointly test whether the coefficients β_2 and β_3 equal zero for the cubic model or similarly for β_2, β_3, d_1, d_2,... d_p equal zero for the spline model. From Table 7 both the cubic and the spline have five regressions rejecting the joint tests, indicating the presence of nonlinearity. The beta plots in Figure 3 also illustrate the nonlinearity.

Next, we evaluate the goodness-of-fit in three measures: $\bar{R}^2$, Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC). We compare the nonlinear test models to the Fama linear model in Table 8. The cubic regression outperforms the Fama version in terms of adjusted $\bar{R}^2$ and AIC: seven $\bar{R}^2$s are better and so are six AICs. The SIC, the most stringent measure, still reports four cubic models that outperform the Fama regressions. The spline regressions also have an advantage in $\bar{R}^2$ and AIC compared to Fama regressions: all $\bar{R}^2$s are better, so are six AICs. Different from the cubic model, the spline model does a bad job in terms of SIC: only three regressions get improved betas. An additional comparison is between these two nonlinear models, presented in the last row of Table 8. The spline is superior to the cubic only in terms of $\bar{R}^2$, but turns out worse in terms of SIC for all the regressions. Generally speaking, the nonlinear models have clearly better goodness-of-fit than the linear Fama model for $\bar{R}^2$ and AIC, but relative to the cubic the the spline adds little extra and does even worse in terms of SIC.

As mentioned before, the spline regression is recommended when we have no strong priors

Table 7: **Fama, Cubic and Spline regressions**

Fama: $\tilde{s}_{t,\Delta} = \alpha + \beta f_{t,\Delta} + \tilde{\epsilon}_{t,\Delta}$

Cubic: $\tilde{s}_{t,\Delta} = \alpha + \beta_1 f_{t,\Delta} + \beta_2 f_{t,\Delta}^2 + \beta_3 f_{t,\Delta}^3 + \tilde{\epsilon}_{t,\Delta}$

Spline: $\tilde{s}_{t,\Delta} = \alpha + \beta(f_{t,\Delta}) f_{t,\Delta} = \alpha + [\beta_1 + \beta_2 f_{t,\Delta} + \beta_3 f_{t,\Delta}^2 + d_1(f_{t,\Delta} - k_1)_+^2 + ... + d_p(f_{t,\Delta} - k_p)_+^2] f_{t,\Delta} + \tilde{\epsilon}_{t,\Delta}$

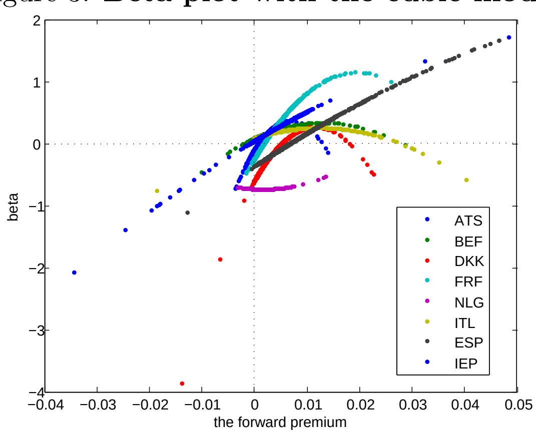
	Fama	Cubic				Spline							
	β	β_1	β_2	β_3	$Wald$	band 1	band 2	band 3	band 4	band 5	band 6	band 7	$Wald$
ATS	*0.11	-0.11	***138.02	-1.00e+4	***4.75	-0.36	0.21	0.01	0.27	-0.12	0.07	***0.31	1.55
BEF	***0.25	0.08	**38.60	-1.47e+3	0.45	***-1.70	***-2.19	*-1.41	-0.08	*0.26	**0.28	***0.27	***31.86
DKK	***0.15	***-0.61	***148.65	-6.35e+3	***22.93	-0.14	0.55	0.49	0.22	***0.47	***0.31	0.11	***142.85
FRF	***0.78	-0.26	***142.65	-3.63e+3	*2.62	***-5.51	-0.13	1.27	*0.73	***0.81	**0.42	***0.95	1.01
NLG	***-0.66	***-0.74	-5.16	1.50e+3	0.08	***-1.97	***-1.54	-0.44	0	-0.31	***-0.46	***-0.65	0.64
ITL	***0.16	0.08	26.30	-1.06e+3	0.75	-0.68	-0.49	-0.03	-0.01	0.23	**0.56	0.04	***4.08
ESP	*0.91	-0.37	55.55	-251.97	*2.38	**3.68	2.73	**1.71	**0.79	**0.56	*0.34	***1.20	***7.89
IEP	*0.15	0.04	***50.40	-329.33	***23.84	-0.96	**-1.22	***-1.79	**-0.36	-0.19	0.23	0.53	***34.68

Key: The Fama panel reports constant Fama βs; the Cubic panel reports all the coefficients of the cubic models and Wald tests for the joint hypothesis that β_2 and β_3 equal zero; the spline panel shows values of $\beta(f_{t,\Delta})$ computed at each band's midpoint value, not the coefficients themselves. The significance tests for each $\beta(f_{t,\Delta})$ are based on the variance-covariance matrix of the coefficients, and the Wald tests show the presence of nonlinearity if the joint hypothesis that all the coefficients of the plus functions and of higher orders of $f_{t,\Delta}$ equal zero.

Table 8: **The Goodness of fit: Fama, Cubic and Spline regressions**

	Fama			Cubic			Spline		
	$\bar{R}^2$	AIC	SIC	$\bar{R}^2$	AIC	SIC	$\bar{R}^2$	AIC	SIC
ATS	0.28%	-8.598	-8.590	1.11%	-8.605	-8.588	1.30%	-8.602	-8.559
BEF	0.96%	-6.899	-6.891	1.06%	-6.898	-6.881	2.37%	-6.907	-6.864
DKK	0.19%	-6.671	-6.663	4.96%	-6.720	-6.701	7.16%	-6.737	-6.694
FRF	6.84%	-6.250	-6.242	7.92%	-6.262	-6.243	8.88%	-6.266	-6.223
NLG	5.13%	-8.131	-8.123	5.03%	-8.129	-8.112	5.62%	-8.130	-8.088
ITL	-0.04%	-4.913	-4.905	0.17%	-4.914	-4.897	0.61%	-4.913	-4.871
ESP	8.46%	-4.799	-4.790	10.90%	-4.842	-4.807	12.19%	-4.834	-4.791
IEP	0.18%	-5.826	-5.187	9.83%	-5.933	-5.912	10.90%	-5.932	-5.884
#(Regressions that outperform Fama)				7	6	4	8	6	3
#(Regressions that outperform Cubic)				—	—	—	8	4	0

Figure 3: **Beta plot with the cubic models**



on the shape of the beta plot. Table 7 reports the estimated mid-point betas of each band. The estimates turn out to be quite imprecise, though; probably as a result of that we do not find any clear pattern. The cubic models, in contrast, produce similar patterns: low or negative values for β_1 (much lower than the standard Fama betas), positive coefficients for the quadratic term in seven cases out of eight, and negative ones for the cubic in again seven cases out of eight. The results are investigated more closely in the following section. At this point we conclude that the Fama model is misspecified because of the presence of the nonlinearity, and is inferior to the nonlinear models in the goodness-of-fit. There is, however, no evidence that we need to go beyond a cubic model. So we now restrict the generalized Fama beta to a simpler form, a simple quadratic function of the forward premium f.

3.2 The Cubic Model: Further Discussion

From Table 7, we see that there is a nonlinear relation: the coefficients β_2 are significant in five cases out of eight and the Wald tests reject the constant beta in five regressions. Even though the cubic coefficient β_3 is never significant individually (t-test), its sign is always negative except for the NLG, suggesting that the parabola of the beta plots opens downward. Consistent with the negative signs of β_3, Figure 3 shows that the Fama betas have an inverse U-shaped pattern in the forward premia. True, the inverse U-shape is not uniform across currencies: we see only a weak convexity for the ESP and the IEP, and the NLG actually has a slightly concave line; but there certainly is no pattern of falling betas in the forward premium that would have supported the Bansal risk premium, nor is there any U or V shaped pattern as proposed by the transaction cost or limit-to-arbitrage hypothesis. The only story that seems to be consistent with the data, in short, is the Fallen-angel hypothesis. In most cases, the max of the inverse U tends to occur at positive forward premium, 1% to 2%, indicating that the bad vibes seem to start not when forward premia fall below zero but already when they fall below a critical positive level.

To control for a possible regimes shift, we now separate the whole sample into two parts, subsample 1, before August 3rd, 1993, and subsample 2 after this date. Under heavy speculation pressure, on August 3rd, 1993 the ERM exchange rate fluctuation bands were widened to $\pm 15\%$ around the central parities. Figure 4 depicts the movements of the forward premia. Compared with subsample 1, the forward premia in subsample 2 move within a much narrower zone. This might mean that the investors expect less speculation with the wide fluctuation band; but towards to the end of the sample period they surely also anticipated more financial integration among ERM countries. Whatever the reason, danger signals occur much more rarely in subsample 2 and the fallen-angel effect should be hard to document.

Panel A of Table 9 summarizes the empirical results from the cubic models for the whole sample and the two subsamples. The slopes in the first subsample are quite close to the slopes in the whole sample. But in the second subsample, there are more egregious numbers, whether in the coefficients β_1 or the higher-order coefficients β_2 and β_3. We find similar results later, when the regressor is the filtered forward premium or the co-movement component: in the post-93 period, the regressors show very little variation, and as a result, the estimated coefficients are erratic.

3.3 The Decomposition of the Forward Premia

Table 9: **The Cubic models with different sample periods and various regressors**

Panel A: regressor is f

$$\tilde{s}_{t,\Delta} = \alpha + \beta_1 f_{t,\Delta} + \beta_2 f_{t,\Delta}^2 + \beta_3 f_{t,\Delta}^3 + \tilde{\epsilon}_{t,\Delta}$$

	whole sample				subsample 1				subsample 2			
	β_1	β_2	β_3	$Wald$	β_1	β_2	β_3	$Wald$	β_1	β_2	β_3	$Wald$
ATS	-0.11	***138.02	-1.00e+4	***4.75	-0.07	135.19	-1.00e+4	***4.74	-0.48	636.92	3.16e+5	0.86
BEF	0.08	**38.60	-1.47e+3	0.45	-0.02	40.73	-1.42e+3	0.46	-2.73	678.63	-3.19e+4	1.35
DKK	***-0.61	***148.65	-6.35e+3	***22.93	-0.84	144.10	-5.67e+3	***21.35	-4.36	23.42	4.89e+4	***93.40
FRF	-0.26	***142.65	-3.63e+3	*2.62	-0.89	201.81	-5.17e+3	**3.82	4.24	-4.19e+3	7.74e+5	1.86
NLG	***-0.74	-5.16	1.50e+3	0.08	***-0.92	8.13	1.51e+3	0.26	-0.67	-276.67	2.20e+5	0.50
ITL	0.08	26.30	-1.06e+3	0.75	1.26	-75.59	1.01e+3	1.98	12.88	-1.78e+3	-1.41e+5	***29.98
ESP	-0.37	55.55	-251.97	*2.38	-0.54	*59.96	-279.16	2.00	12.00	-4.00e+3	2.97e+5	***15.06
IEP	0.04	***50.40	-329.33	***23.84	-0.06	**47.52	-236.02	***21.91	4.37	-4.19e+3	8.51e+5	***6.69

Panel B: regressor is ff

$$\tilde{s}_{t,\Delta} = \alpha + \beta_1 ff_{t,\Delta} + \beta_2 ff_{t,\Delta}^2 + \beta_3 ff_{t,\Delta}^3 + \tilde{\epsilon}_{t,\Delta}$$

	whole sample				subsample 1				subsample 2			
	β_1	β_2	β_3	$Wald$	β_1	β_2	β_3	$Wald$	β_1	β_2	β_3	$Wald$
ATS	0.25	117.42	-1.26e+4	***7.73	0.26	122.86	-1.31e+4	***8.19	-0.60	260.80	3.36e+5	0.42
BEF	-0.15	68.16	-1.99e+3	*2.42	-0.03	59.66	-1.86e+3	1.50	*-1.98	469.89	-1.28e+4	2.09
DKK	-0.49	***124.63	-6.24e+3	***30.65	-0.38	111.82	-6.23e+3	***21.06	***-3.69	-88.36	6.42e+4	***88.00
FRF	-0.64	**441.00	-1.87e+4	***18.64	-0.57	*441.92	-1.73e+4	***15.84	-4.02	-2.01e+3	1.59e+6	***4.70
NLG	***-1.27	-68.98	-349.88	1.17	***-1.27	62.11	253.75	1.07	-0.88	-661.89	3.07e+5	1.81
ITL	**-1.20	110.14	-1.55e+3	*2.93	-1.18	372.77	-1.51e+4	***6.33	***-1.91	-3.89e+3	-1.85e+5	***37.13
ESP	*0.36	**85.81	422.63	***4.21	**0.45	**76.67	562.51	**3.54	-5.65	-3.33e+3	9.29e+5	***18.07
IEP	0.05	**49.37	-302.77	***23.82	0.19	49.84	-338.40	***25.36	-1.91	-3.89e+3	-1.85e+5	***37.13

Panel C: regressor is $\widehat{f}$

$$\tilde{s}_{t,\Delta} = \alpha + \beta_1 \widehat{f}_{t,\Delta} + \beta_2 \widehat{f}_{t,\Delta}^2 + \beta_3 \widehat{f}_{t,\Delta}^3 + \tilde{\epsilon}_{t,\Delta}$$

	whole sample				subsample 1				subsample 2			
	β_1	β_2	β_3	$Wald$	β_1	β_2	β_3	$Wald$	β_1	β_2	β_3	$Wald$
ATS	-1.27	1.82e+3	-5.23e+5	*2.58	-1.13	1.68e+3	-4.85e+5	1.27	-0.30	-1.98e+4	-1.03e+8	1.48
BEF	-1.20	854.09	-1.00e+5	***12.06	**-3.49	1.56e+3	-1.60e+5	***14.39	***36.50	***-9.32e+4	***-6.87e+7	**3.89
DKK	-2.25	**684.55	**-4.38e+4	**3.01	***-5.32	1.26e+3	-7.56e+4	***6.49	***112.51	***-9.45e+4	***2.51e+7	***8.79
FRF	-2.39	*909.82	-6.20e+4	**3.09	**-4.31	***1.34e+3	**-9.14e+4	**4.49	36.28	1.55e+4	-3.07e+6	***6.08
NLG	-1.16	***4.26e+3	***-1.84e+6	***6.46	-1.58	**5.08e+3	***2.15e+6	***6.91	**-33.71	**-2.14e+5	-3.78e+8	***5.43
ITL	***-11.56	***1.56e+3	-5.71e+4	***11.76	***-15.22	***1.91e+3	***-6.75e+4	***9.00	-389.22	1.66e+5	-2.31e+7	***7.57
ESP	***-10.75	***1.37e+3	-4.61e+4	***11.97	***-14.63	***1.76e+3	***-5.83e+4	***12.91	**-232.38	**9.86e+4	-1.35e+7	***8.25
IEP	***-4.98	***1.48e+3	***-1.10e+5	***4.90	***-8.93	***2.30e+3	***-1.63e+5	***9.09	***85.78	***-1.13e+5	***4.20e+7	***5.04

Figure 4: **The movement of the forward premium**

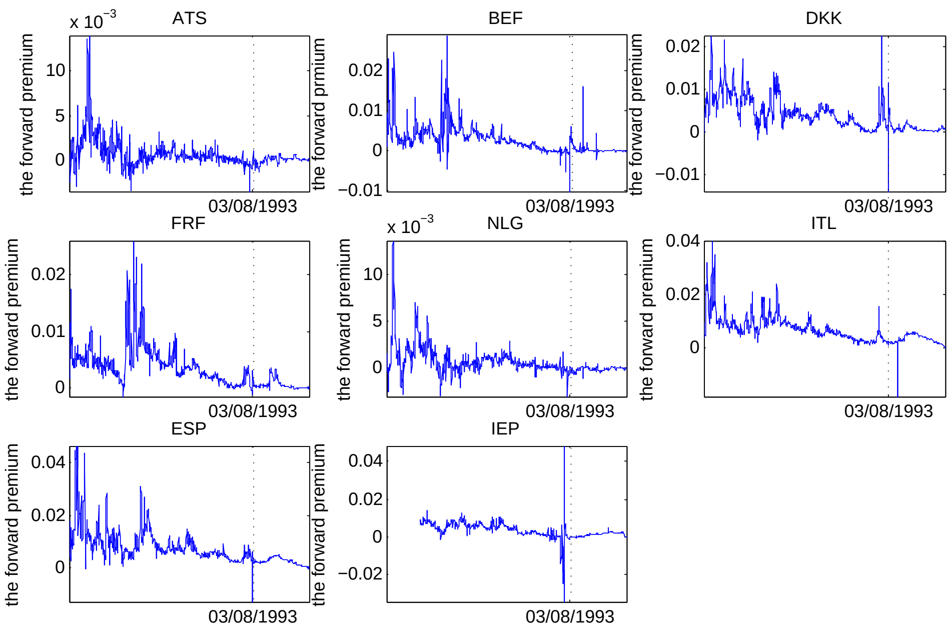


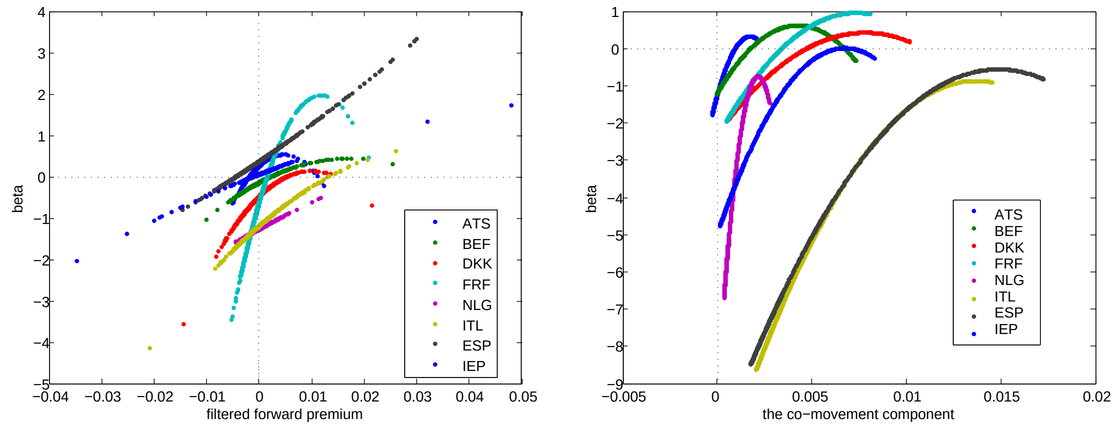
Table 10: **Comparing the filtered premium or the common trend to the original forward premium in terms of goodness of fit.**

Number of times the alternative right-hand-side variable beats the original one in terms of fit.

	$R^2_{ff\succ f}$	$AIC_{ff\succ f}$	$SIC_{ff\succ f}$	$R^2_{\widehat{f}\succ f}$	$AIC_{\widehat{f}\succ f}$	$SIC_{\widehat{f}\succ f}$
whole sample	7	7	6	2	2	2
subsample 1	7	7	7	2	2	2
subsample 2	5	6	6	4	5	5

Recall that the nonstationary forward premium can be decomposed into a long-memory co-movement component $\widehat{f}$ and a short-term filtered forward premium ff as described in Section 2. The common component $\widehat{f}$ is nonstationary and is a good candidate for the long-memory part of the risk premium or for slow-changing features in the base country's monetary policy or some other economic fundamentals. Whatever the underlying, the slow-moving component $\widehat{f}$ is unlikely to explain the short-term expectation on the forward contracts. In addition, its nonstationary nature creates statistical problems in the regressions. In that sense, $\widehat{f}$ might be responsible for the forward puzzle. The filtered component, ff, in contrast, is stationary and may be closer to the expectations, but it may of course also pick up any transient component of a risk premium. Panel B and Panel C in Table 9 test the cubic models where we use each of these two components as the regressor instead of their sum, the forward premium.

Figure 5: **Beta plots in terms of the components**



(a) Beta plot with the filtered forward premium

(b) Beta plot for the co-movement component

Compared with the raw f in Panel A, the filtered forward premium ff produces higher coefficients β_1 and more significant Wald tests in the first subsample. In addition, ff produces less egregious slopes in the second subsample. For the whole sample, four regressions have higher coefficients β_1 with the ff, and more regressions (seven out of eight) reject the linear model specification via the significant Wald test. The co-movement component, in contrast, brings out much lower and negative coefficients in the whole sample and subsample 1, and many erratic coefficients β_1 in subsample 2. (In the second subsample, the lack of variability in the forward premia mentioned above may also contribute to make the betas even worse.) The more negative β_1s are consistent with the idea that the co-movement component is related to the long-memory part of the risk premium, responsible for the missing variable(s); meanwhile, the more positive β_1s with ff are consistent with the conjecture that the filtered forward premium ff loads more heavily on the expectation than $\widehat{f}$. The cubic slope β_3 stays negative in most cases, demonstrating an inverse U-shaped pattern of the beta in Figure 5(a) and Figure 5(b). The inverse U-shaped pattern of beta plots in ff again supports the Fallen-angel hypothesis that the career risk premium of the portfolio managers becomes crucial to the investment decision when the danger signals appear in the markets. But the pattern is far less pronounced than for the raw f. With $\widehat{f}$ as the regressor, in contrast, the pattern is very clear, whether statistically or graphically.

Table 10 assesses the goodness-of-fit — $\bar{R}^2$, AIC and SIC — when the cubic model is run on different specifications of the forward premia. All three measures indicate a better fit for the filtered forward premia than for the raw forward premia: seven out of eight regressions show an improvement in the whole sample and the first subsample; six regressions gain in terms of AIC and SIC in the second subsample. In contrast, the co-movement component does worse than the original regressor with six regressions worsening in the whole sample and subsample 1, and with mixed results in subsample 2. So, the model evaluation is consistent with our conjecture that the short-term filtered forward premium loads on the expectation or on the transient part of the risk premium possibly, having higher predictive power; while, the long-memory co-movement deteriorates the forward prediction, is responsible for the missing variables.

4 Conclusions

This paper wants to sort out two issues in the forward-puzzle literature: the model-misspecification problem when a non-linearity is not recognized, and the issue of a non-stationary forward premium. The latter creates a statistical issue (how does one discover the correct error margins?) and one of economic interpretation: how can an expectation be non-stationary, especially for ERM currencies?

One idea we bring into this literature is career risk considerations among money managers: to go against publicly visible danger signals takes either guts or extra expected returns. In an exploratory test we find that danger signals other than the forward premium do seem to be followed by extra expected excess returns. The next question is whether the forward premium may also double as a danger signal, in which case the Fama beta should be allowed to be a quadratic function or a spline function of the forward premium. We do find that a nonlinear relationship is more proper to describe the relation between the percentage change in spot rate and the forward premium. Both the cubic and spline regressions outperform the linear model in terms of goodness-of-fit. The cubic model produces an inverse U-shaped beta that rises from very negative to a higher level as the forward premium approaches zero and dips again as the forward premium increases further. This pattern again fits in with the Fallen-angel hypothesis that traders or portfolio managers shun long positions in assets with danger signals.

The second idea we explore is to decompose the highly persistent forward premium into a long-memory co-movement component and a filtered forward premium that is definitely not unit-root. The former component associates with a strong common trend. It seems to load heavily on the missing regressor which, in Fama's conjecture, is responsible for the forward puzzle: when we run the cubic models with the co-movement variable on the right-hand side, betas are much lower than for the total-f regressions and the filtered-ff regressions. The shape of the betas even provides a clue as to what the long-memory part stands for: it could again be a Fallen-angel risk premium that makes traders skeptical about assets with danger signals. The filtered forward premia, in contrast, seems to load on both the expectations and the risk premium: the betas are generally positive and do a better job indicating the future change, but also exhibit a (weaker) inverse U-shaped pattern that one expects from career-risk effects.

References

Albuquerque, R., 2006, The Forward Premium Puzzle in a Model of Imperfect Information: Theory and Evidence, working paper, University of Rochester.

Alex, M., 2005, Testing forward rate unbiasedness allowing for persistent regressors, *Journal of Empirical Finance* **12**, 613-628.

Ardian, H. and Wade Brorsen, B., 2002, the Overlapping Data Problem, working paper, SSRN.

Backus, D. and Smith G., 1993, Consumption and Real Exchange Rates in Dynamic Economies with Non-traded Goods, *Journal of International Economics* **35**, 297-316.

Backus, D., Foresi, S. and Telmer, C., 2001, Affine term structure models and the forward premium anomaly, *Journal of Finance* **56**, 279-304.

Baillie, R. T., and Kiliç, 2006, Do Asymmetric and Nonlinear Adjustments Explain the Forward Premium Anomaly?, *Journal of International Money and Finance* **25**, 22-47.

Bansal, R., 1997, An Exploration of the Forward Premium Puzzle in Currency Market, *the Review of Financial Studies* **10**, 369-403.

Bilson, J.F.O., 1981, The Speculative Efficiency Hypothesis, *Journal of Business* **54**, 435-451.

Chinn, M. and Frankel, J., 2002, Survey Data on Exchange Rate Expectations: More Currencies, More Horizons, More Tests, *Monetary Policy, Capital Flows and Financial Market Developments in the Era of Financial Globalizsation: Essays in Honour of Max Fry (London: Routledge)*, 145-167.

Cumby, R. and Obstfeld, M., 1984, International Interest Rate and Price Level Linkages under Flexible Exchange Rates: A Review of Recent Evidence, in J. Bilson and R. Marston eds. *Exchange Rate Theory and Practice*, University of Chicago Press, Chicago, IL.

Cumby, R.E., 1987, Is it Risk? Explaining Deviations from Uncovered Interest Parity, *Journal of Monetary Economics* **22**, 279-299.

Domowitz, I. and Hakkio, C.S., 1985, Conditional Variance and the Risk Premium in the Foreign Exchange Market, *Journal of International Economics* **19**, 47-66.

Engel, C., 1996, The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence, *Journal of Empirical Finance* **3**, 123-192.

Fama, E., 1984, Forward and Spot Exchange Rates, *Journal of Monetary Economics* **14**, 319-338.

Frankel, J. and Engel, C., 1984, Do Asset-demand Functions Opitimize Over the Mean and Variance of Real Return? A Six-currency test, *Journal of International Economics* **17 (3-4)**, 309-323.

Froot, K. and Thaler,R., 1990, Anomalies: Foreign Exchange, *Journal of Economic Perspectives* **4**, 179-192.

Hodrick, R. J. and Srivastava,S., 1986, The Covariation of Risk Premiums and Expected Future Spot Exchange Rates, *Journal of International Money and Finance* **5**, 5-22.

Hodrick, R. J., 1988, The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Market, *Journal of International Economics* **25(3-4)**, 387-390.

Hodrick, R. J., 1989, Risk, Uncertainty and Exchange Rates, *Journal of Monetary Economics* **23**, 433-459.

Hollifield, B., and Uppal, R., 1997, An Examination of Uncovered Interest Parity in Segmented International Commodity Markets, *Journal of Finance* **52**, 2145-2170.

Huisman, R., Koedijk, Kee., Kool, C. and Nissen, F., 1998, Extreme support for uncovered interest parity, *Journal of International Money and Finance* **17**, 211-228.

Ikenberry, D., Lakonishok, J., and Vermaelen, T., 1995, Market Underreaction to Open Market Share Repurchases, *Journal of Financial Economics* **39**, 181-208.

John S.R. et al., 1979, A Probability Distribution and its Uses in Fitting Data, *Technometrics* **21(2)**, 201-214.

Lyons, R. K., 2001, The Microstructure Approach to Exchange Rates, MIT Press.

Mark, N.C., 1988, Time-varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts, *Journal of Financial Economics* **22**, 335-354.

Ramberg, R., Dudewicz, E.J., Tadikamalla, P.R., Mykytka, E.F., 1979, A probability Distribution and its Uses in Fitting Data, *TECHNOMETRICS*, **21(2)**, 201-214.

Ramsay, J.O., and Silverman, B.W., 2005, Functional Data Analysis, *Springer.*

Obstfeld, M. and Rogoff, K., 2000, New Directions for Stochastic Open Economy Models, *Journal of International Economics* **50(1)**, 117-153.

Roll, R. and Yan, Sh., 2000, An Explanation of the Forward Premium Puzzle, *European Financial Management* **6(2)**, 121-148.

Sarno, L. Valente, G. and Leon, H., 2006, Nonlinearity in Deviations from Uncovered Interest Parity: An Explanation of the Forward Bias Puzzle, *CEPR Discussion Papers*, 5527.

Scott, W., Barnhart and Andrew C.S., 1991, Testing the Unbiased Forward Rate Hypothesis: Evidence on Unit Roots, Co-Integration, and Stochastic Coefficients, *Journal of Financial and Quantitiative Analysis* **26(2)**, 245-267.

Sercu, P. and Vinaimont, T., 2006, The Forward Bias in the ECU: Peso Risks *v.* Fads & Fashions, *Journal of Banking and Finance*, **30**, 2409-2432.

Wayne, E.F., Sergei, S. and Timothy, S., 2002, Spurious Regressions in Financial Economics?, NBER working paper 9143.

Takagi, Sh., 1989, An Alternative Approach to the Divergence Indicator in the European Monetary System, *DE ECONOMIST*, **137(3)**, 360-366.

Tkacz, G., 2001, "Estimating the Fractional Order of Integration of Interest Rates Using a Wavelet OLS Estimator", *Studies in Nonlinear Dynamics and Econometrics*, **5 (1)**, Article 2.

Wolff, C., 1987, "Forward Foreign Exchange Rates, Expected Spot Rates and Premia: A Signal-Extraction Approach", *Journal of Finance*, **42(2)**, pp. 395-406.

A Appendix: Monte-Carlo Standard Deviations

The remaining issue is the reliability (SE) of the estimations. There are two complications. First, the forward premium (and even more so the co-movement component) is non-stationary or nearly so. Second, following Hansen and Hodrick (1980) we do not want to waste information by considering only non-overlapping forward contracts, so we use the weekly observations on one-month forward contracts. For either reason, the conventional standard deviation is underestimated. In this paper Monte Carlo (MC) simulation is employed to calculate the standard deviation.

Monte Carlo method is a stochastic technique that uses random numbers and probability statistics rather than math to discover the distribution of parameter estimates. Variables on both sides of the classical UEH model are expressed in their Moving Average (MA) models. According to the correlogram test, we find that for most currencies the first six lags are significant,[11] so the exchange rate change and the forward premium are expressed as AR(6) models:

$$s_{j,t} = \alpha_1 + \beta_1 s_{j,t-1} + \beta_2 s_{j,t-2} + ... + \beta_6 s_{j,t-6} + \nu_{j,t}, \tag{1.21}$$

$$f_{j,t} = \alpha_2 + \theta_1 f_{j,t-1} + \theta_2 f_{j,t-2} + ... + \theta_6 f_{j,t-6} + \xi_{j,t}. \tag{1.22}$$

It turns out that the residuals ν and ξ are non-normally distributed. Edward and John (1979) provide a technique for a non-normal distribution number generator. This technique accommodates a broad class of distributions because it transforms a uniform random number into distribution with any desired set of values for the first four statistical moments (mean, variance, skewness and kurtosis). These four moments, denoted below as μ_1, μ_2, μ_3 and μ_4, are functions of four parameters $\lambda_1, \lambda_2, \lambda_3$ and λ_4, as described in the following equations:

$$\mu_1 = \lambda_1 + \frac{A}{\lambda_2}, \tag{1.23}$$

$$\mu_2 = \frac{B - A^2}{\lambda_2^2}, \tag{1.24}$$

$$\mu_3 = \frac{C - 3AB + 2A^3}{\lambda_2^3}, \tag{1.25}$$

$$\mu_4 = \frac{D - 4AC + 6A^2B - 3A^4}{\lambda_2^4}. \tag{1.26}$$

In these equations, the terms A, B, C and D are also functions of $\lambda_1, \lambda_2, \lambda_3$ and λ_4:

$$A = \frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4}, \tag{1.27}$$

$$B = \frac{1}{1+2\lambda_3} + \frac{1}{1+2\lambda_4} - 2\,\mathcal{B}(1+\lambda_3, 1+\lambda_4), \tag{1.28}$$

$$C = \frac{1}{1+3\lambda_3} - \frac{1}{1+3\lambda_4} - 3\,\mathcal{B}(1+2\lambda_3, 1+\lambda_4) + 3\,\mathcal{B}(1+\lambda_3, 1+2\lambda_4), \tag{1.29}$$

$$D = \frac{1}{1+4\lambda_3} + \frac{1}{1+4\lambda_4} - 4\,\mathcal{B}(1+3\lambda_3, 1+\lambda_4) + 6\,\mathcal{B}(1+2\lambda_3, 1+2\lambda_4) - 4\,\mathcal{B}(1+\lambda_3, 1+3\lambda_4), \tag{1.30}$$

[11] Remember that these are overlapping weekly observations of one-month returns and premia.

where $\mathcal{B}(u, v)$ is the beta function. To generate the residuals we estimate their first four moments and numerically solve for the corresponding values of the λ's. The desired non-normal random number $\tilde{R}$ is the following transformation of a unit uniform random number $\tilde{p}$:

$$R(\tilde{p}; \lambda) = \lambda_1 + \frac{\tilde{p}^{\lambda_3} - (1 - \tilde{p})^{\lambda_4}}{\lambda_2}. \tag{1.31}$$

To test the null of no relation between s and f we generate 1000 numbers with the properties of the observed s_j, and, independently of that, 1000 numbers with the properties of the observed f_j. (The actual number of observations in the real-world sample is between 1030 and 1200.) We then run on this data set all regressions we study in this paper and compute all betas we are interested in. We repeat this 1000 times to get an idea of how the regression output is distributed under the null.